

Does talent migration increase inequality? A quantitative assessment in football labour market

Vasilakis, Chrysovalantis

Journal of Economic Dynamics and Control

Published: 31/12/2017

Peer reviewed version

[Cyswllt i'r cyhoeddiad / Link to publication](#)

Dyfyniad o'r fersiwn a gyhoeddwyd / Citation for published version (APA):

Vasilakis, C. (2017). Does talent migration increase inequality? A quantitative assessment in football labour market. *Journal of Economic Dynamics and Control*, 85(December), 150-166. <https://dial.uclouvain.be/pr/boreal/en/object/boreal%3A196664>

Hawliau Cyffredinol / General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Does Talent Migration increase Inequality? A quantitative assessment in Football Labour Market

Chrysovalantis Vasilakis *

Abstract

I analyze the links between talent migration and cross-country inequality by exploiting the 1995 elimination of mobility restrictions on the European football labor market. I develop a simple model and employ an empirical dataset to estimate its parameters. Through simulation analysis, I compare actual data with a counterfactual no-mobility restriction trajectory, and conclude that the elimination of mobility barriers increases not only cross-country inequality by 25%, but also global output in the football economy by stimulating the production of new talent in Africa, Latin and Central America.

Keywords: International Migration, Brain Drain, free mobility, Inequality, European Football.

JEL Codes: F22, J61

*Bangor Business School E-mail: chvasilakis@gmail.com

1 Introduction

The market for talent is becoming increasingly globalized. From 1990 to 2000, the stock of high-skilled migrants to 30 OECD countries rose from 12.5 to 20 million (Docquier, Lowell, and Marfouk, 2009). As such, there is greater propensity for elite workers, such as engineers, physicians, and researchers, to migrate (Docquier and Rapoport, 2012). However, the main issue is how globalization affects inequality. The answer to this question is complex, as the link between migration and inequality is bi-directional (de la Croix and Docquier, 2012). Income differentials induce migration, which affects cross-country inequality. In turn, migration affects cross-country inequality because migrants tend to be highly skilled, as opposed to workers who do not migrate, both domestically and in foreign countries. Furthermore, lack of availability of long-term longitudinal data on migration size and structure, as well as the absence of quantitative data on migration policies, makes it difficult to identify the mechanisms at work.

This paper is a step toward investigating the effects of talent migration on inequality across countries by focusing on the European football labor market. The football market offers four significant advantages for the study of mobility and inequality. First, professional football is a small but very visible segment of the highly skilled labor market, as a significant proportion of the world's population closely follows this sport. As a result, any significant change in the set of rules governing football hierarchy is of interest per se.

Second, professional football has one of the most globalized labor markets. Since the mid-1970s, an increasing number of talented players have moved across countries and continents to (re-)join the top European leagues. By 2010, about 50% of the world's most talented players are employed in a foreign country. This percentage is substantial compared to the global average emigration rate of college-educated workers, which is only around 6% (Docquier, Lowell and Marfouk, 2009). Consequently, similar to Kleven et al.'s (2013) analysis, the patterns observed in the football market are likely to provide an upper limit for the migration inequality response to the elimination of mobility restrictions for highly skilled workers and its impact on the global economy.

Third, a unique panel dataset on European male footballers' mobility since 1974 can be exploited. This allows identifying the most talented players by country of origin, documenting international mobility, and building metrics for leagues and national team performance. Limited theoretical models have studied the interactions between the brain-drain and global inequality. Although existing models are all calibrated to match data, it is still unclear whether the implemented specifications are the most appropriate, or if their micro-foundations are well supported empirically (see de la Croix and Docquier, 2012; Docquier and Rapoport, 2012). This paper contributes to this literature by developing an endogenous migration decisions model exploiting data features and estimating its structural parameters based on country-specific characteristics.

Fourth, the crucial advantage of analyzing the football market is that a significant and unex-

pected policy reform was implemented in 1995, drastically changing the migration patterns of talented players. This reform, known as the Bosman rule, considerably eased player transfers by eliminating transfer fees for out-of-contract players wishing to change clubs within the European Union (EU) and banning the inclusion of EU citizens in "foreign player" quotas. Consequently, mobility of both European and non-European players across European clubs increased.

The use of a quasi-natural experiment, i.e., the Bosman rule, allows quantifying the effect of a liberalization of migration shocks on cross-country inequality. Specifically, I develop a random utility model for the football players' decision to migrate to certain leagues, where each player's utility depends on both the characteristics of the targeted foreign league and moving cost. I also assume a production technology for the output of each European league and national team, relating the relative quality of each country to the number of talented players and their average skill level. To the best of my knowledge, this is the first model that characterizes the migration technology of talented football players in a panel setting.

My empirical estimations show the following. On one hand, the decision to migrate for a talented player is affected by common variables, such as income per capita in both country of origin and destination, geographical distance, and colonial links. On the other hand, one of the key factors is the prestige/quality of the destination league. Furthermore, I show that the total number of talented players is a key determinant of average league performance, and the total number of talented players weighted by the score of their employment league is a key determinant of national team performance.

Moreover, my structural model also allows quantifying the impact of the Bosman rule on talent training and production, and I find that it spurs these in poorer regions, such as Africa, Central and South America (except in Argentina and Brazil). However, I do not identify such incentive effects in European or Asian countries. This final finding corroborates the brain-gain theory (Beine et al., 2011; Docquier and Rapoport, 2012) according to which skill-biased emigration prospects affect national talent production in poorer countries.

Subsequently, I conduct counterfactual simulations to assess Bosman rule's effect on player mobility, their geographic concentration, efficiency, and inequality of leagues and national teams. I show that the Bosman rule led to an improvement in world football quality as measured by Union of European Football Associations (UEFA) and Federation Internationale de Football Association (FIFA) scores. European football average quality improved by 12% in 2010, and the national teams' average quality by 20%. This result originated from the incentive training mechanism developed by the introduction of the Bosman rule as discussed in the counterfactual analysis. Moreover, the Bosman rule increased inequality across European leagues and decreased inequality across national teams. On one hand, the latter effect seems to be mainly caused by the removal of mobility barriers, which may have reinforced the concentration of best talents in top European leagues and, thus, amplified the difference in quality between the top and the rest of the European leagues. On the other hand, migrant players may have improved their skills by playing

in higher quality leagues. Consequently, this may have helped their respective national teams to perform better. These effects are sizeable; the Gini index of UEFA scores shows an increase of 25% in cross-country inequality between European leagues from 1998 to 2010, whereas cross-country inequality between national teams fell by 70% on average during the same period.

This paper is also linked to the sport economics literature that focused on the effects of globalization of European soccer on game quality and national teams' inequality like Milanovic (2005), who use data from World Cup performances, the Champions League, and Italian Series A League. He finds that, consistent with his model, as the market globalized (i) the concentration of quality for teams participating in the Champions League had increased steadily since 1983, (ii) rich clubs from north Italy displaced poor clubs from the south starting in the 1990s, and (iii) inequality in national teams' performance in World Cups had decreased starting in the 1980s. Frick (2009) explores the effects of Bosman rule in migration and its effects to the national teams. Radoman (2017) analyzes the impact of Bosman ruling, on the productivity of new entrants to the English Premier League. My study contributes to this literature not only through quantifying the Bosman rule's effects in inequality and efficiency of the national teams and European leagues through the counterfactual analysis but also through showing the Bosman rule's impact on the production of new talents across different regions, which is the major source of efficiency gain.

The remainder of this paper is organized as follows: Section 2 presents the data development, Section 3 describes the theoretical model, and Section 4 its parametrization and illustrates a quasi-experiment. Finally, I present my conclusions in Section 5.

2 Data development

In this section, I describe the database, which includes the location of talent, bilateral migration stocks, the performance of national leagues and national teams, and other data sources used to estimate the framework's equations.

A good indicator of an athlete's talent is if he is considered "the best" in his country of origin, while simultaneously being recognized not only by football professionals but also by the general public. The definition of talent in literature varies. A widely-used definition of "talented player" considers players among the 23 selected to represent their national teams in the FIFA competition, according to FIFA regulations (see Frick, 2009; Berlinschi et al., 2013). For instance Berlinschi et al. (2013) consider a talented player as one who participates in national teams for two different years even if the team's composition changed over time. Additionally, Frick (2009) defines talent by only considering the 23 players selected into the national team for competition. Another strand of literature defines "talented players" as those who play in major leagues (see Kleven et al., 2013 and Radoman, 2017). Furthermore, the sports literature defines "talent" as a highly valued athlete based on team supporters' preferences (see Rosen, 1981).

In this paper, I define a talented player as one who has at least three appearances in one of the

best 65 national teams during a World Cup year. This definition has the advantage to incorporate the definition of talent consistency for outstanding performances. Talented players are those who "regularly" represent their national teams and not simply participate once.

My analysis focuses on the nine World Cup years, from 1978 to 2010 (the World Cup takes place every four years). Talented football players make more appearances in these years because (i) trainers use their best players, (ii) there are more matches played for the World Cup qualification and final stages, and (iii) all highly ranked football-playing nations play a similar number of games. I select the best 65 countries based on the 2011 FIFA rankings. These include 21 UEFA core countries, 29 non-EU countries (mainly from Africa and Latin America), and 15 Eastern European countries. The latter group includes 12 new countries created after 1990 (three of them were former countries: ex-USSR [split in 1991]; ex-Yugoslavia [split in 1991]), and ex-Czechoslovakia [split in 1993]).

The utilized key variables are as follows:

- I denote by N_{ijt} the number of talented players (talents) originating from the national team i and employed in league j during World Cup year t . The data are provided by Benjamin Starck-Zimmermann and the CIES Football Observatory, which records the names of players appearing in their national team, their number of appearances, and their club of employment on an annual basis. The total number of talents originating from country i is given by $N_{it}^T = \sum_j N_{ijt}$, whereas the total number of talents employed in league j is denoted by $l_{jt} = \sum_i N_{ijt}$.
- I denote by Q_{it} the quality/score of national team i in year t . The data are provided by the official FIFA website and start from 1994.¹ However, the rules do change over the years. To make results comparable, I standardize scores, dividing the FIFA score of each country in my dataset by Spain's score, $Spin$, for the same year. I choose Spain because it is the last World Cup winner in my sample. This normalization makes the variable more stationary and comparable over time, and has no incidence on the result.
- I denote by q_{jt} the quality/score of league j in year t . I measure the league's quality by its UEFA score. The data are available from 1978 to 2010, but only for the European leagues.² To make the results comparable, I standardize the league score. For the national teams' scores, I divide each league's annual reported UEFA score by Spain's score. I choose the Spanish league as the basis of normalization because Spain is the best European league in the 2010 UEFA ranking. UEFA's score is based on the total number of points accumulated in European club competitions by clubs in each league over the past five years. For each

¹See <http://fr.fifa.com/worldranking/rankingtable/index.html>. This measure is constructed from FIFA to capture the quality of the national teams and it is used extensively in the literature (Ericson, 2000; Berlinschi et al. 2013).

²There can be alternative measures of league quality such as the number of tickets "sold". I find that the correlation between the number of tickets with the league's performance is high. In a sample of seven leagues from 1978 to 2010, I conclude that the correlation between the two variables was around 0.77.

country, it is computed as the sum of all points obtained by participating teams and then divided by the number of clubs. Points are assigned according to specific parameters, such as 1) the number of games played by a team, 2) the number of teams qualified in second and following rounds, and 3) the number of goals scored by a team in the entire competition.³

Other variables I also use a number of variables to estimate the model. Data on geographic distance, colonial links, and linguistic proximity are taken from the CEPII database (see Clair et al., 2004). UEFA authorities provided data on average wages. Specifically, I collected average wages for 21 EU leagues from 1978 to 2010 from their UEFA archives. To include leagues outside Europe in the regressions, I input missing salaries by using a regression of average salary and income per capita. As shown below, this is a reasonable strategy, since the correlation between these two variables is particularly high.⁴ This strategy allows me to avoid large scale differences when using income per capita as a proxy of wages for the missing salaries. I collect the data for income per capita from World Bank indicators. The wages are deflated by the national price indices. As a result, I use real wages in my analysis. Additionally, I include a proxy for institutional quality from Polity IV, which covers the period from 1978 to 2010 for the 65 countries in my sample. Furthermore, the marginal tax rates for soccer players are provided by Kleven et al. (2013) and are limited only to 14 EU countries for the same period. As Kleven et al. (2013) claim, considering the top marginal tax rates has the advantage to be uncorrelated with salary levels. Finally, I calculate the training variable as follows: I record the number of stadiums and training centers for each of the 65 countries/leagues in the sample from 1978 to 2010. I construct this variable using the official websites of each country's club and using Wikipedia to find the year the stadium was built, used, closed, or re-opened.

3 Football Economy

Quantifying the Bosman rule effect requires modeling interdependencies between migration and football economy performances. I first develop a simple micro-founded model that highlights how the mechanisms of underlying migration decisions and inequality characterized the football market's function. I also endogenize the football players' training decisions, which govern the

³Data for the quality index of leagues are only available from 1978 onwards, and data on q_{jt} only for EU leagues. To proxy the score of non-EU leagues, I adopt the following procedure: I regress the European national teams' scores on the quality/score of their league and identify a linear relationship. I subsequently proxy the quality/score of the league of non-EU countries by using this linear equation. Hence, I have two samples: the 1978–2010 balanced group of EU-to-EU countries and another group that includes both the balanced group and the 1978 to 2010 group of all-to-EU countries. EU-to-EU means that the country of origin is in the EU as is the destination. On the other hand, all-to-EU represents the country of origin for players from all over the world with a EU destination. The linear combination between the leagues' and national teams' scores is $q_{i,t} = 0.6397Q_{i,t} - 0.0237$. The regression explains 30% of the variation in the league's quality/score. The average correlation between the leagues' and national teams' scores between 1994 and 2010 is 0.58, but it gradually increases over time. This proxy allows me to increase the sample for the migration equation and provide a robustness analysis.

⁴The correlation between the two variables is 0.87, and the regression is $\log \text{wage} = 47 \cdot \log \text{GDP} - 5000$.

supply of talent.⁵

3.1 Migration Decisions

Many studies investigate the size and structure determinants of labor mobility between countries. To endogenize migration decisions, most recent research is built on a multinomial discrete choice model without spatial correlation in unobserved factors (Beine et al., 2011; Grogger and Hanson, 2011; Kennan and Walker, 2011; Canaday and Tamura, 2009). I use similar micro-foundations to analyze the determinants of talented players' mobility.

Each country i produces a number N_{it}^T of talented players with heterogeneous preferences. The utility of a player from country i deciding to play in league j is the sum of a deterministic component, depending on the destination league's characteristics and an individual-specific random term, which captures the soccer players' heterogeneous attitudes and incentives. Possible determinants include the average wage level (w_{jt}), the prestige of the league (measured by its standardized UEFA score, q_{jt}), a vector a_{jt} of amenities (capturing quality of life, the size of the total diaspora, league rules, quality of stadium and infrastructure, etc.), and mobility costs $c_{ijt} \geq 0$ if league j differs from country i (it is assumed that $c_{iit} = 0$). I assume that the deterministic utility component is logarithmic. I also assume that the random term ε_{ijt} follows an iid extreme-value distribution and it has the property of independence from irrelevant alternatives (IIA). Thus, it is independent of the observed attributes. I omit player subscripts for clarity.

The utility of a player from country i who remains and plays in league i at time t is given by

$$u_{iit} = \theta \ln w_{it} + \delta \ln q_{it} + \rho \ln a_{it} + \varepsilon_{iit},$$

where (θ, δ, ρ) is a vector of preference parameters.

The utility for the same player moving to league j at time t accounts for moving costs (c_{ijt}). These costs capture psychological and integration costs, monetary costs (which can be partly covered by the player's team of employment), as well as screening costs (the effort of contacting foreign employers and obtaining a visa) and quotas.⁶

The utility of moving to league j is given by

$$u_{ijt} = \theta \ln w_{jt} + \delta \ln q_{jt} + \rho \ln a_{jt} + \ln(1 - c_{ijt}) + \varepsilon_{ijt},$$

where c_{ijt} denotes bilateral migration cost, modeled here as an "iceberg" cost, and it includes

⁵In the Web Appendix, I analyze the mechanisms of underlying migration decisions and inequality in a two-league setting, and I consider the Bosman rule's impact on the football economy.

⁶Some federations (especially the European leagues) impose restrictions on the employment of foreign players. For instance, a football player must fulfill two conditions to play in the English league: he must play 75% of his national team's matches and his national team's rank must be greater than the 70th FIFA position. Football clubs in Norway need to demonstrate that having a foreign player would improve the league and club's level. The Spanish league does not allow more than three non-EU players per club.

all the various types of costs previously discussed. I use the McFadden theorem (McFadden, 1984), which is when the random term follows an iid extreme-value distribution, where the probability that a talented soccer player born in country i chooses to emigrate to league j follows a logit expression:

$$\begin{aligned}\frac{N_{ijt}}{N_{it}^T} &= \Pr[u_{ijt} = \text{Max}_k u_{ikt}] \\ &= \frac{\exp[\theta \ln w_{jt} + \delta \ln q_{jt} + \rho \ln a_{jt} + \ln(1 - c_{ijt})]}{\sum_k \exp[\theta \ln w_{kt} + \delta \ln q_{kt} + \rho \ln a_{kt} + \ln(1 - c_{ikt})]}.\end{aligned}$$

Hence, the ratio of emigrants to stayers is given by the following expression:

$$\begin{aligned}\frac{N_{ijt}}{N_{iit}} &= \frac{\exp[\theta \ln w_{jt} + \delta \ln q_{jt} + \rho \ln a_{jt} + \ln(1 - c_{ijt})]}{\exp[\theta \ln w_{it} + \delta \ln q_{it} + \rho \ln a_{it}]} \\ &= (1 - c_{ij,t}) \left(\frac{w_{jt}}{w_{it}}\right)^\theta \left(\frac{q_{jt}}{q_{it}}\right)^\delta \left(\frac{a_{jt}}{a_{it}}\right)^\rho.\end{aligned}\quad (1)$$

The bilateral emigration rate (ratio of emigrants to stayers) is a function of differentials in wages, prestige levels, amenities, and net migration costs.

3.2 Technologies

3.2.1 Production Function of Leagues

Another strand of literature examines the impact of immigration on economic performance and welfare in destination countries (see recent works by Borjas, 2003; Card, 2009; Docquier et al., 2015). Similarly, I use a simplified version of their models to build the technology of leagues. Each league has many homogeneous clubs characterized by the same production function. The quality/score in league j (q_{jt}) is

$$q_{jt} = A_{jt} l_{jt}^\gamma, \quad (2)$$

where A_{jt} is the scale productivity factor in league j in year t , $l_{jt} \equiv \sum_i N_{ijt}$ is the total number of talented players employed in the league, and γ is the elasticity of the leagues' score to the number of talented players.

3.2.2 Production Function of National Teams

Each nation has a national team that participates in the different FIFA competitions. Each national team is composed by the most talented players in the country, regardless of the league in which they employed. Talented players improve their skills when playing in better leagues. For simplicity, the talented player's skill is assumed to be proportional to the average quality of the

league in which he is employed. I denote by $L_{it} \equiv \sum_j N_{ijt} q_{ijt}$ the total amount of skills available to the national team i at time t . Berlinschi et al. (2013) construct a similar index. They weight each migrant with the ranking of the league in which employed. The performance of this national team is depicted by the following technology:

$$Q_{it} = B_{it} L_{it}^{\phi}, \quad (3)$$

where B_{it} is the scale productivity factor in country i in year t , and ϕ is the elasticity of the national teams' scores to the total amount of skills used by the national team.

3.3 Endogenous Training decisions

Recent literature shows that skill-based migration prospects affect expected returns to schooling and, thus, incentives to acquire education (see Mountford, 1997; Beine et al., 2011; Theodoridis, 2014). However, the direction is ambiguous. On one hand, individuals have more incentive to invest ex-ante in education due to higher returns to schooling. On the other hand, some of them migrate, thus decreasing the average number of skilled workers in the country of origin. There is a net brain gain once the first effect dominates.

In the same line, I assess whether the elimination of mobility barriers-the Bosman rule-induces more football players to train harder and become more talented. The Bosman rule has changed football players' migration patterns, by reducing migration hurdles and increasing talent "exportability." As a result, it may have increased in the players' incentives to acquire skills. Many growth models with human capital assume that human capital during period $t + 1$ is described by a Cobb-Douglas function, which depends on the level of human capital and on the level of investment in schooling at a given time t (see Lucas, 1988; Azariadis and Drazen, 1990, Tamura, 1992; Tamura, 1994). In a similar vein, I assume the following talent production:

$$N_{it+1}^T = Z_{it} (N_{it}^T)^{\sigma} \exp(\xi_{it}), \quad (4)$$

where Z_{it} is a scale factor, characterizing the training technology of country i in year t ; σ is a technological parameter capturing inertia in the number of talents (one period is four years and a career lasts longer than four years); and ξ_{it} is the proportion of young players (aged, e.g., 16 to 20) who train to become talented professional players at time $t + 1$ (the exponential form is used for mathematical simplicity).

Becoming a professional football player requires training, but extra effort $\ln E$ is needed to become a talented professional football player. This level of effort is heterogeneous across young players and uniformly distributed on $[0, 1]$. Players decide how much effort to put into training without knowing their future migration inclinations (i.e., the level of individual-specific random term ϵ_{jt+1} , which is an extreme value distributed in the ex-post utility function). Based on de

Palma and Kilani (2007), who show that under the Type I Extreme Value distribution, the unconditional and conditional distributions of maximum utility coincide, I assume that the expected utility function of becoming a talented player is given by:

$$u_{it+1}^e = \ln \sum_j e^{[lnv_{ijt} + \ln(1-c_{ijt})]} - \ln E,$$

where $lnv_{ijt} = \theta lnw_{jt} + \delta lnq_{jt} + \rho lna_{jt}$. I also denote the expected utility of a "regular" (non-talented) player as $v_{t+1}^e = lnv_{ot}$. I assume that regular players stay in their home countries. Players decide to train if $\ln \sum_j e^{[lnv_{ijt} + \ln(1-c_{ijt})]} - \ln E > lnv_{ot}$. It becomes straightforward to show that the proportion of players training harder to be considered talented is given by:

$$\xi_{it} = \sum_j \frac{v_{jt}}{v_o} (1 - c_{ijt}). \quad (5)$$

The proportion depends on the magnitude and distribution of bilateral migration costs; as bilateral migration costs decrease, the expected return to training increases, and more players invest in training. Substituting (5) into (4) and taking the logs yields:

$$\ln(N_{it+1}^T) = \ln Z_{it} + \sigma \ln N_{it}^T + \xi_{it} \quad (6)$$

$$= \ln Z_{it} + \sigma \ln N_{it}^T + \frac{v_{it}}{v_o} + \sum_{j \neq i} \frac{v_{jt}}{v_o} (1 - c_{ijt}) \quad (7)$$

Finally, I consider the following equilibrium for my "football economy":

Definition 1. For a given distribution of talented players, $\{N_0^T\}_{\forall i}$, and set of exogenous variables, $\{w_{it}, a_{it}, A_{it}, B_{it}\}_{\forall it}$ and $\{c_{ijt}\}_{\forall ijt}$, an inter-temporal equilibrium is a set of year-specific vectors $\{N_{it}^T, q_{it}, Q_{it}\}_{\forall it}$ and matrices $\{N_{ijt}\}_{\forall ijt}$ satisfying migration optimality conditions (1), technological constraints (2) and (3), training optimality conditions (7) and aggregation constraints $l_{it} \equiv \sum_j N_{jit}$ and $L_{it} \equiv \sum_j N_{ijt} q_{ijt}$.

4 Quantitative Analysis and parametrization

This section presents the identification strategy and empirical estimation for each equation of the theoretical framework as set out in the previous section. By estimating each equation, I find elasticities using all sources of variation in migration decisions, quality/score of leagues, national teams' scores, and the number of talents in all leagues over time. Subsequently, I conduct a counterfactual analysis and quantify the effects of the Bosman rule in the inequality and efficiency of leagues and national teams.

4.1 Identification

4.1.1 Migration Decisions

In this subsection, I develop the identification strategy to estimate the Eq. (1), which presents the migration decisions of soccer players. The Bosman rule affects migration costs; consequently, for simplicity, I consider that \widehat{c}_{ijt} denotes the pre-Bosman level of migration costs and the Bosman rule multiplied $(1 - c_{ijt})$ by a constant. Therefore, I write the migration's cost of Eq. (1) as $(1 - c_{ijt}) = (1 - \widehat{c}_{ijt}) \cdot \exp(\lambda BOS_t)$, where λ is a parameter and BOS_t is a variable equal to zero in the pre-Bosman period and equal to one after the Bosman rule. Then, Eq. (1) can be expressed as:

$$\frac{N_{ijt}}{N_{iit}} = (1 - \widehat{c}_{ijt}) \exp(\lambda BOS_t) \left(\frac{w_{jt}}{w_{it}} \right)^\theta \left(\frac{q_{jt}}{q_{it}} \right)^\delta \left(\frac{a_{jt}}{a_{it}} \right)^\rho. \quad (8)$$

Particularly, using data on the bilateral migration of players (N_{ijt}) and real wage levels for leagues (w_{jt}), measuring the league's prestige (q_{jt}) by its standardized UEFA score, I estimate the parameters of the structural Eq. (8). The migration costs are not observable. Following empirical literature on migration, I allow migration costs to increase with the log of the geographic distance between countries (d_{ij}), and decrease with linguistic proximity (l_{ij}) and with the existence of colonial links between countries (s_{ij}). I also control for polity ($polity_{jt}$) at destination to capture the possible quality of institutions at destination as an attraction for soccer players. The intuition is that several developing countries face political instability and, thus, football players might emigrate to a country/league with better institutions. Additionally, I control for marginal tax rates at destination. Finally, migration costs are affected by the Bosman rule, BOS_t . Since 1995, European football players have been allowed to move without any legal restriction within Europe, and restrictions on non-EU players have been attenuated. Hence, I include a Bosman dummy (BOS_t) in the set of regressors.

Taking the log of the Eq. (8), I obtain the following pseudo-gravity specification:

$$\begin{aligned} \ln \frac{N_{ijt}}{N_{iit}} = & \theta \ln \frac{w_{jt}}{w_{it}} + \delta \ln \frac{q_{jt}}{q_{it}} + \lambda BOS_t \\ & + \beta_1 \ln d_{ij} + \beta_2 l_{ij} + \beta_3 s_{ij} + \psi polity_{jt} + \mu \ln(1 - tax_{jt}) + \alpha_i^m + \alpha_j^m + f(t) + \epsilon_{i,t}^m, \end{aligned} \quad (9)$$

where $(\lambda, \beta_1, \beta_2, \beta_3)$ are parameters governing the migration cost equation, (α_i^m, α_j^m) are fixed effects for origin and destination countries, $f(t)$ is a B-spline function that captures the nonlinear time trend, and $\epsilon_{i,t}^m$ is the error term typically assumed to be statistically independent of regressors. The fixed effects, together with the distance and Bosman variables, control for unobserved characteristics and migration costs. The main coefficients of interest are δ and λ . Other explanatory variables are time invariant and exogenous. I cannot estimate Eq.(9) by including

both time-fixed effects and the Bosman dummy because of multi-collinearity. To identify this dummy variable's effect on migration flows, I estimate Eq. (9) using a non-linear time trend by a polynomial of t , $f(t)$, the B-spline method, which substitutes the year-fixed effects approach. This nonparametric method captures the common year-effects before and after the adoption of the Bosman rule, as the institutional changes in the champion's league occurring in 1992, the increase in TV rights and wages, which increased in all the leagues after 2003 (see Kessene, 2009).

I also include the marginal tax rates in Eq. (9) to control different taxation policies for soccer players or foreign workers, which alternated in a few countries before and after 1995 (see Kleven et al., 2013), such as the Belgian reform in 2002, the Danish reform in 1992, and the Spanish reform in 2004.⁷ The correlation between real wages and marginal tax rates was small (0.02), ruling out a risk of collinearity between these two variables when I included both in Eq. (9).

As far as estimation is concerned, the OLS-FE regression technique is likely to yield inconsistent estimates of the coefficients because of the presence of a high proportion of zero values (83%) in bilateral migration stocks (N_{ijt}). Estimating the gravity model (9) with OLS-FE led to the exclusion of numerous country pairs. If the country pairs with zero flows have a different population distribution from pairs with positive flows, this exclusion generates a selection bias. Additionally, as it is clearly identified by Santos-Silva and Tenreyro (2006), it is possible the variance of $\epsilon_{i,t}^m$ to depend on its covariates of $\frac{N_{ijt}}{N_{iit}}$ and then, its expected value will also depend on some of the regressors in the presence of zeros. This, in turn, invalidates one important assumption of consistency of OLS estimates. The use of the Poisson pseudo-maximum likelihood (PPML) estimator, allows us to include the zero values for the dependent variable, and rules out any selection bias of this kind. Therefore, I use Poisson regressions as in other studies on trade and bilateral migration (Beine et al., 2008, 2011; Tenreyro, 2007). Moreover, the standard errors of Eq.(9) are clustered by country of origin and destination.

Finally, estimation of Eq.(9) may also be subject to a reverse causality problem. The reason is that, while migration has an impact on the league's quality, it also likely depends on the league's prestige. In practice, it is not obvious that this problem is severe. The total immigration and emigration affect league quality. On average, each bilateral pair accounted for a relatively small fraction of total migration stocks. Nevertheless, I address the reverse-causality problem by using IV-Poisson and instrumenting the log ratio of the leagues' scores by their lagged value.

⁷Since 2002, international football and basketball players in Belgium (playing in either first or second league) have the option of paying a flat income tax rate of 18% instead of the regular progressive income tax schedule with very high rates at the top. Conversely, in 1992, Denmark enacted a special tax scheme for foreign researchers and high-income foreigners in all other professions, who signed employment contracts in Denmark after June 1, 1991. The Spanish reform is known as "Beckham Law." The "Beckham Law" (Royal Decree 687/2005) is a particular tax scheme passed in 2005, applicable to foreign workers having moved to Spain after January 1, 2004.

4.1.2 League Production Function.

To estimate the parameters of the production function Eq. (2), I use standardized data on league scores and data on the distribution of talented players. I model scale productivity factors by using country- and time-fixed effects, which capture the individual league characteristics and the common changes that happen to all leagues, such as reforms in the fundamental form of European competitions. Taking the log of Eq. (2) yields:

$$\ln q_{jt} = \alpha_j^q + \alpha_t^q + \gamma \ln l_{jt} + \epsilon_{jt}^q, \quad (10)$$

where (α_j^q, α_t^q) are the vectors of fixed effects and ϵ_{jt}^q is the error term. First, I estimate this equation by OLS-FE. As previously explained, OLS might be subject to a reverse-causality bias. Indeed, the league's quality impacts migration decisions and the total number of immigrants and emigrants. I address this problem by using IV regressions and employ two different instruments. Similarly to Beine et al., (2008), I use firstly the lagged number of talented players. Secondly, following Feyrer (2009), Beine and Parsons, (2015) and Docquier et al. (2016) I instrument migration stocks using predicted migration stocks obtained from a "zero stage," pseudo-gravity regression that includes interactions between year dummies and the geographic distance between each country of origin and each destination. The pseudo-gravity model is:

$$\ln N_{ijt} = \beta_t \ln d_{ij} + \nu X_{ij} + a_i + a_j + a_t + \epsilon_{i,t}, \quad (11)$$

where the vector X_{ij} is a set of controls, such as colonial relationship and linguistic proximity, and a_i, a_j, a_t are the origin, destination, and year fixed effects. In the pseudo-gravity stage, the high prevalence of zero values in bilateral migration stocks gives rise to econometric concerns about possible inconsistent OLS estimates. To address this problem, I use the Poisson regression by pseudo-maximum likelihood (see Santos Silva and Tenreiro, 2006).⁸

4.1.3 National Production Function

To estimate the parameters of the production function Eq.(3), I use standardized data on national teams' scores and on the distribution of talented players. I model scale productivity factors by using country- and time-fixed effects. I obtain the following empirical specification by taking the log of the Eq. (3):

$$\ln Q_{jt} = \alpha_j^Q + \alpha_t^Q + \phi \ln L_{jt} + \epsilon_{jt}^Q, \quad (12)$$

where (α_j^Q, α_t^Q) are the vectors of fixed effects and ϵ_{jt}^Q is the error term. Initially, Eq. (12) was estimated by OLS-FE. I address the reverse causality problem by using IV regressions and employ

⁸In the Web appendix, I explain the methodology of constructing this instrument.

two different instruments. The first instrument is the lagged variable of $\ln L_{jt}$. The second one, used for robustness, are the predicted migration stocks obtained from a zero-stage pseudo-gravity regression as in Beine and Parsons, (2015), and Docquier et al. (2016), weighted for the league's score.

4.1.4 Training Equation

I develop the empirical model for the training equation from the micro-founded condition (Eq. (7)) and, since the Bosman rule decreases migration costs, I consider that $(1 - c_{ijt}) = (1 - \hat{c}_{ijt}) \cdot \exp(\lambda BOS_t)$ as in the migration equation. Then, Eq. (7) becomes

$$\ln(N_{it+1}^T) = \ln Z_{it} + \sigma \ln N_{it}^T + \frac{v_{it}}{v_o} + \sum_{j \neq i} \frac{v_{jt}}{v_o} (1 - \hat{c}_{ijt}) \exp(\lambda BOS_t) \quad (13)$$

and $C_{it}^1 = \frac{v_{it}}{v_o}$ and $C_{it}^2 = \frac{v_{jt}}{v_o} (1 - \hat{c}_{ijt}) \cdot \exp(\lambda)$, where C_{it}^2 could be proxied by regional dummies, so that

$$\ln(N_{it+1}^T) = \ln Z_{it} + \sigma \ln N_{it}^T + C_{it}^1 + C_{it}^2 BOS_t \quad (14)$$

Since I use a beta-convergence specification (i.e., I subtract $\ln N_{it}^T$ from the right- and left-hand sides and use $\beta_0 = \sigma - 1$ in Eq. (15)). This convergence regression model is in line with Beine et al.'s (2008) specification. I include the log change in the total number of country-specific talented players on its lagged value and a Bosman dummy in the regression model. I allow the Bosman effect to vary across different regions. Specifically, I distinguish Europe, Africa, Central and South America, and Brazil and Argentina. Brazil and Argentina (referred to as BA), the two countries outside of Europe that have strong leagues and national teams before the Bosman rule, are distinguished from the group of other Latin American countries. I account for unobserved heterogeneity by adding a full set of country- and time-fixed effects and some controls, such as training ($train_{it}$) and polity ($polity_{it}$). I include the last two variables to capture the countries' possible training investments and governance quality:

$$\Delta \ln N_{it}^T = \alpha_i^N + \alpha_t^N + \beta_0 \ln N_{it-1}^T + \sum_r \varphi_r BOS_t d_{ir} + \theta_1 polity_{it} + \theta_2 \ln train_{it} + \epsilon_{it}^N \quad (15)$$

where Δ is a time difference operator ($\Delta x_t = x_t - x_{t-1}$), (α_i^N, α_t^N) is the vector of fixed effects, β_0 is the convergence parameter expected to lie between -1 and 0, d_{ir} is a dummy equal to one if country i belongs to region r and 0 otherwise, and φ_r is a coefficient capturing the incentive effect induced by the Bosman rule in region r . My set of control variables includes institutional quality and the leagues and clubs' training expenditures. As a benchmark, I estimate this equation using OLS-FE. The OLS regression model assumes that all covariates are independent of the error term. Although fixed effects control for possible misspecification caused by unobserved characteristics,

they do not account for other potential sources of regressor endogeneity.

Endogeneity arises because of the presence of the lagged term on the right-hand side. The use of fixed effects and AR terms lead to estimation inconsistency (Nickell, 1981). As such, it is preferable to use two-stage least-squares (2SLS) methods and instruments $\ln N_{i,t-1}$, with their lagged values. Furthermore, as Caselli et al. (1996) argue, empirical studies on convergence are plagued by the incorrect treatment of country-fixed effects. It is usually assumed that these effects are uncorrelated with the other right-side variables. The fixed effects α_i^N is used as a determinant of the log-change number of talented players. By construction, it is also a determinant of the lagged term, which is the regressor in the equation. Hence, the assumption of uncorrelated fixed effects is violated in dynamic panel regressions. It is desirable to correct for this collinearity bias even if I use fixed effects as control variables.

To solve this problem, I use Arellano and Bond's (1991) approach to developing a generalized method of moments (GMM) estimator using all moment conditions. This method suggests eliminating country-fixed effects using differences. The equation can be rewritten as:

$$\Delta \ln N_{it}^T = \bar{\alpha}_t^N + (1 + \beta_0) \Delta \ln N_{i,t-1}^T + \sum_r \varphi_r \Delta BOS_t d_{ir} + \theta_1 \Delta polity_{it} + \theta_2 \Delta \ln train_{it} + \bar{\epsilon}_{it}^N, \quad (16)$$

where $\bar{\alpha}_t^N \equiv \alpha_t^N - \alpha_{t-1}^N$ is the new time-fixed effect, $\bar{\epsilon}_{it}^N \equiv \epsilon_{it}^N - \epsilon_{it-1}^N$ is the transformed error term, and $\Delta BOS_t = BOS_t - BOS_{t-1}$.

4.2 Results

4.2.1 Migration decisions.

Table 1 presents the results of estimating Eq. (9) using different samples and estimation methods. Columns (1)-(3) report the results of the main sample, which it contains a set of origin and destination leagues of 21 European core countries from 1978 to 2010. Columns (4)-(5) highlight the results of estimating Eq. (9) for a larger sample covering the 1978 to 1994 EU-to-EU countries and the 1994-2010 All-to-EU countries. Also, Columns (1)-(2) and Columns (4)-(5) present the Poisson approach and Column (3) reports the IV-Poisson method results.

Most regressions explain more than 50% of the variation in the observed migration flows, thus ensuring an excellent fit, given that the migration rates are very heterogeneous across corridors. The main explanatory variables' coefficients are significant and have intuitive signs. The coefficients are stable across estimations and different samples. Precisely, the prestige ratio variable is significant. Moreover, its coefficient is positive, and lies between 0.43 and 1.31 (see Columns (1) to (5)). The latter implies that a 1% increase in the prestige ratio led to an increase of 0.4% to 1.31% in the migration flows between the two countries. Similarly, the wage ratio's coefficient lies between 0.82 and 1.2. Conversely, the larger the geographic distance between origin and destination, the

lower the migration flows are. Moreover, the language variable is not significant in most regressions, probably because the soccer industry overcomes linguistic problems. Football teams either hire translators or allow players to communicate in their native language. Additionally, colonial links are not significant, inferring that culture similarities are not necessary for the soccer player's decision to migrate.

Furthermore, I introduce the taxation variable in Columns (2) to (5). The sample decreases because tax data are only available for 14 destination countries instead of the 21 in Column (1). The sign of the taxation coefficient is negative in most regressions, and not significant (Columns (2) to (5). Columns (1) to (3) show that polity is not significant, but it is positive and significant at 5% in Columns (4) and (5). The explanation for this result is that Columns (1) to (3) include only European countries of origin, while Columns (4) and (5) include countries from all over the world as origin. The intuition behind this result can be that a soccer player originating from an EU country could face similar institutional quality across EU countries. Conversely, for a football player who originates from a developing country, better institutions can improve his living conditions.

The Bosman dummy is one of the most important variables. It is significant and positive. The coefficient is 0.7 on average in all regressions, which means that the Bosman rule leads to an increase of 7% of migration flows on average. It is important to observe that this coefficient varies across the different sub-samples, lies between 0.48 to 1, and remains significant at 5%.

Furthermore, when analyzing the large sample covering the 1978 to 1994 EU-to-EU countries and the 1994 to 2010 all-to-EU countries, I split the Bosman dummy into two dummies: Bosman EU-to-EU ($Bosman_{EU}$) and the Bosman non-EU-to-EU ($Bosman_{NonEU}$), to investigate whether the Bosman rule creates differential effects depending on the soccer player's country of origin. Column 5 shows that both dummies are significant and positive. The coefficient of $Bosman_{NonEU}$ is slightly higher than the coefficient of $Bosman_{EU}$ (0.88 versus 0.62). A Wald test on the difference of coefficients between $Bosman_{EU}$ and $Bosman_{NonEU}$ dummies (Column 5) shows that the two coefficients are not statistically different.⁹

The last issue to be explored is the investigation of the existence of reverse causality between the prestige ratio and migration flows. Comparing Poisson and IV-Poisson regressions, the Wald test shows that there is no significant evidence for reverse causality. The explanation of no endogeneity can be because prestige q_{jt} depends on the sum of N_{ijt} ($\sum_i N_{ijt}$), which was the total number of talented players employed in the league, as defined in the Eq. (2). This means that each N_{ijt} of Eq. (9) accounted, on average, for 1/65th of the $\sum_i N_{ijt}$. Thus, endogeneity is small to be detected and the Wald test is not significant.

[Insert Table 1]

⁹In the Web appendix, I provide a falsification test generating fake Bosman between 1986 and 1998. Then, I include each fake Bosman dummy in Eq. (9) to investigate whether it affects migration decisions. I find that the Bosman rule was an unexpected event.

4.2.2 League Production Function.

Table 2 reports the estimation results of Eq. (10). All regressions include full sets of country-fixed effects and year dummies. Moreover, I control for heteroscedasticity and serial correlation, and I report the clustered standard errors in parentheses. Column (1) provides OLS-FE results for two different groups. Column (1) includes all European leagues, and Column (2) includes European and Eastern European leagues. Columns (3) to (5) provide the results of the IV method using different groups. Column (3) illustrated the results of the full group. For robustness, Column (4) contains the estimation results with only the 21 European leagues, excluding countries from the ex-USSR and ex-Yugoslavia. In the end, Column (5) includes all European leagues. For Columns (3) to (5), I use the 2SLS estimation technique with the lags of $\ln l_{jt}$ as internal instruments and exogenous instruments as per Feyer (2009) (see Column 5) as I explain in the identification strategy of league production.

Most regressions explain more than 69% of the variation in observed league performance. In all variants, there is a significant and positive impact on the number of talented players on the league's performance. Precisely, if the number of migrants increases by 1%, the league's performance increases by 0.54% on average. This coefficient is less than 1 and shows that the technology is a concave transformation. Thus, there are decreasing returns to scale in the production function. Furthermore, the Hausman test shows evidence of endogeneity. Consequently, IV regressions are preferable to OLS-FE. Additionally, my regressions fulfill the two necessary conditions for instrumentation. Since Cragg-Donald cannot be strictly valid in the presence of heteroscedasticity, I use the "rule of thumb" of an F-stat above 10 to test for the presence of weak instruments. In all first stage regressions, F-stats are always above 10 in Columns (2) to (5). In other words, not all possible instruments (internal and external) are weak. Moreover, the Cumby-Huizinga test shows no evidence of serial correlation for the regressions that uses lags as instruments (see Columns (3)–(4)).

4.2.3 National Team Production Function

Table 3 illustrates the results of the estimation of Eq. (12). It shows that the total quantity of talent in efficiency units (i.e., the total number of talented players weighted by the performance of the league of employment) is significant and positive in all columns under different samples and methodologies. In Column (1), the coefficient was very low (0.08) using OLS-FE. I also use 2SLS to solve the endogeneity problem in Eq. (12). The first two lags of $\ln L_{it}$ are used as instruments to solve the endogeneity problem (see Column (3)). I also use the external instrument of Feyrer (2009) to solve the endogeneity issue in a different way (see Column (2)). The Hausman test shows evidence of endogeneity in Columns (2) and (3). Consequently, the IV regression is preferable to OLS-FE. The two necessary conditions for instrumentation are thus fulfilled in the regressions. The elasticity in the IV regressions increases on average to 0.3 compared to the OLS-FE estimation.

The Cumby-Huizinga test shows that there is no evidence of serial correlation when I use lags as instruments (see Column (3)).

[Insert Table 3]

4.2.4 Training Equation

Table 4 reports the results of OLS, IV, and Arellano-Bond (1991) regressions for the main variables of interest. Columns (1) and (2) show that the lagged level had a significant and negative impact on the growth rate of talented players. This is in line with Column (3), which shows that the absolute value of $1 + \beta_0$ is lower than 1, implying that the training process is stable and predicts monotonic (conditional) convergence of the number of talented players in the long run. For OLS regression, which is illustrated in Column (1), the speed of convergence is around 0.69. As a result, it takes around 1 period to reach long-run equilibrium. Column (3) shows the Arellano-Bond results. The coefficient of the lagged term, $\Delta \ln N_{i,t-1}$, is around 0.43 and highly significant. Therefore, the speed of convergence is about 0.59 per period. It takes 1.17 periods to reach the country-specific steady state when the explanatory variables are kept constant.

In all regressions, the effect of the Bosman rule on the growth rate of talented players in Africa, Central and South America is positive and significant. In these regions, the incentives to acquire skills have increased since the globalization shock. On the contrary, no such incentive effect is evident for European and Asian countries. Moreover, the incentive effect for the BA group (Argentina and Brazil) is negative but not significant. The explanation behind this finding could be that the market of Argentinian and Brazilian players was already liquid before Bosman rule and so a young player in these countries did not see his incentives change much after this rule. The elasticities are robust to the specification and estimation method. The GMM estimation in Column (3) gives the most accurate estimated coefficients. As observed, the largest incentive effect was in Latin American countries (elasticity of 0.70), followed by Africa and Central America (elasticities of 0.66 and 0.59, respectively). Precisely, the number of talented players in South American countries (except Brazil and Argentina) has increased by two times because of the Bosman rule. I use the lags for the number of talented players as instruments and, in the last column, I instrument training using its lag. The two necessary conditions for instrumentation were fulfilled by the regression in Column (2). Since the Cragg-Donald test is not strictly valid in the presence of heteroscedasticity, I use the "rule of thumb" of an F-stat above 10 to test for the presence of weak instruments. This condition is fulfilled, so the instruments could be considered strong. At the bottom of Column (3), the test of second-order autocorrelation in the residuals shows that there is no evidence of additional serial correlation. Furthermore, the Hansen J-test shows that over-identification restrictions are not rejected. The polity variable is not significant in columns, and the training variable has a positive coefficient and it is significant only in Column 3.

4.3 Parameters and Algorithm

It is necessary to provide a quantitative analysis to identify and quantify the effects of the Bosman rule on inequality and efficiency. The main reason to perform this analysis is that previous estimations show that the Bosman rule has the differential impact on training incentives across nations. For instance, it did not spur training in countries in Europe, Asia, Argentina, and Brazil. Additionally, the magnitude of the incentive effect varies across the other regions. Moreover, most countries are both sending emigrants and receiving immigrants, which lead the Bosman rule effect to have ambiguous consequences for cross-country inequality and efficiency.

Calibration. The main coefficients of interest for the migration equation are the elasticity of migration to prestige, δ , and the effect of the Bosman dummy, λ . The estimate for δ is around 0.7, and those for λ from 0.65 to 1. I use the value of 0.7 for δ and the average value of 0.77 for λ . I use a single value of the Bosman dummy for all the countries of origin because the Wald test shows that there is no significant difference between the $Bosman_{EU}$ and $Bosman_{nonEU}$ dummies. I also use a value of 0.53 for γ , the elasticity of the leagues' score to the number of talented players (Column (5) in Table 2), and a value of 0.2 for ϕ , the elasticity of national teams' score to the total amount of skills (Column 3 in Table 3). For the training equation, I employ the estimated coefficients from Column 3 in Table 4. In particular, I set the $1 + \beta_0$ for the lagged number of talented players at 0.41 and the region-specific coefficients for the incentive effect, φ , at 0.66 for Africa, 0.71 for South America, 0.59 for North and Central America, and 0 for the other regions. Parameter values are summarized in Table 5.

Simulation Algorithm. The algorithm's structure is as follows: first, the training trajectory is simulated as if the Bosman rule has never been implemented; subsequently, the equations for interdependencies between migration decisions and leagues' scores 1998 to 2010 are solved. To do this, I build a Gauss-Seidel "shooting" algorithm: each iteration I starts with a set of 309 guesses, $\tilde{q}_{it}^I \forall i, t$, for the leagues' scores. I plug these guesses into the migration equation to simulate the counterfactual allocation of talented players across countries. I then use these inputs for the production functions and compute the solution for the leagues' scores, $\bar{q}_{ij,t}^I$. I iterate these until the solution coincide with the guesses: the next iteration $I + 1$ starts with a new set of guesses, $\tilde{q}_{it}^{I+1} = \eta \tilde{q}_{it}^I + (1 - \eta) \bar{q}_{it}^I$, where $1 - \eta$ is the correction factor. I use $\eta = 0.95$. The algorithm stops when the sum of errors (in absolute values) falls below a convergence threshold: $\sum_{i,t} |\tilde{q}_{it}^I - \bar{q}_{it}^I| < \epsilon$.

[INSERT TABLE 5]

4.4 Impact of Bosman rule in Inequality and Efficiency

In this subsection, I analyze the Bosman rule's effects on inequality, efficiency, and scores of European leagues, and compare the trajectory with those obtained from the simulation of the situation without the adoption of the Bosman rule. Moreover, I present the Bosman rule's effects on in-

equality, efficiency, and scores of national teams.¹⁰

Efficiency and Inequality for European Leagues. First, I investigate the Bosman rule's overall effect on the efficiency and inequality indices computed for the 23 European leagues. Figure 1 shows that the Gini and Theil indices increased around 25% and 33%, respectively, from 1998 to 2010. Figure 1 also presents the Bosman rule's effect on inequality (Gini index) and efficiency (average score of European leagues) from 1994 to 2010. I report the difference between the observed and the "no-Bosman" trajectories as percentages of the no-Bosman values between 1994 and 2010. It shows that the Bosman rule has led to a sharp short-run increase in the Gini index (+24% in 1998) and a slightly larger long-run effect. In 2010, the Gini index was 28% greater than in the no-Bosman case. I conclude that the Bosman rule explained the entire rise in inequality observed in European football over the past 15 years. It is also apparent that the Bosman rule increases European soccer's efficiency. However, the effect is gradual. The European leagues' average score increases by 3% in 1998 and by 8% in 2010. The figure illustrates that efficiency increases at a lower rate over the 1998 to 2010 period. Additionally, the efficiency of the total number of talented players increased by 15% in 1998 and by 22% in 2010. This globalization shock leads to an improvement in the quality of European football and increases disparities between major and minor leagues.

[Insert Figure 1].

Efficiency and Inequality for National Teams. The results for efficiency (average score of national teams) and inequality (Gini index) are presented in Figure 2, which shows the effect of the Bosman rule from 1994 to 2010. The differences between the observed and no-Bosman trajectories are presented as percentages of the no-Bosman values between 1994 and 2010. Similar to the leagues, I observe that the national teams' efficiency increased after the Bosman rule. The rise is sharp in 1998 (+15%) and subsequently slower until 2010 (+20% during the total period).

Contrary to the results obtained for the leagues, the Bosman rule reduces inequality between national teams. There is a dramatic drop in the Gini index of around 15% in 1998. I observe the largest reduction in 2002 (45%), confirming it with the 2002 World Cup results. Almost half of the teams in the last round of 16 teams are non-EU: South Korea, Japan, Senegal, Paraguay, and Mexico. It is the first time that this phenomenon occurred in the World Cup's history. Non-EU countries performed well during the 2010 World Cup held in South Africa, but to a lesser extent.

[Insert Figure 2]

¹⁰In the Web appendix, I provide a similar analysis of the effects of the Bosman rule's effects on the rankings of European leagues and national teams.

4.5 Robustness Analysis

4.5.1 Exogenous and Endogenous Number of Players

Hitherto, I considered the effects of the Bosman rule in training decisions in my football economy. I define the model incorporating endogenous training decisions as the benchmark model. For robustness analysis, I test one alternative model for the quantitative experiment: the baseline model with an exogenous number of talents, which I compare with the benchmark model. I also define two scenarios to examine the simulation results as follows.

- Scenario 1: Benchmark Model
- Scenario 2: Benchmark Model with an Exogenous Number of Talents, $\varphi=0$ of Eq. (15).

Efficiency and Inequality of European Leagues. Figures 3a and 3b present the European league's efficiency and inequality from 1994 to 2010 in the two scenarios defined above. In Figure 3b, I notice that inequality and efficiency increase over time in all scenarios. The efficiency increases less in Scenario 2 than in 1. Consequently, I infer that the greater production of talent in developing countries and their emigration to Europe are key factors explaining the rise in European football efficiency/quality. Regarding the inequality between European leagues, Figure 3a shows that inequality increases similarly in all scenarios. This result suggests that the incentive mechanisms play a minor role in explaining the rising inequality.

[Insert Figure 3]

Efficiency and Inequality of National Teams. Figures 4a and 4b illustrate the national teams' efficiency and inequality from 1994 to 2010. For the analysis of European leagues, I compare efficiency and inequality between the two scenarios. The inequality falls at a slower pace in Scenario 2 than in 1 (see Figure 4a). I thus conclude that the incentive to train in poor regions and the migration of these talented players in high-quality leagues are the main reasons for the sharp decrease in inequality across the national teams. Note that there is an increase in efficiency after the Bosman rule's introduction in all scenarios. Figures 4a and 4b together show that my results and conclusions for inequality and efficiency in the benchmark model are robust.

[Insert Figure 4]

5 Conclusions

Complex interactions between highly skilled migration and aggregate performance govern the dynamics of growth and inequality across nations. Due to the lack of data, the economic and migration literature have insufficiently studied these interdependencies. I benefit from the availability of a unique data set on European football and the unexpected elimination of mobility

restrictions (the Bosman rule) in 1995 to analyze the links between talent migration and cross-country inequality. I develop a simple model and estimate its parameters. Then, I conduct a simulation analysis, compare the actual data to a counterfactual no-mobility restriction trajectory. The elimination of mobility barriers increases both cross-country inequality by 25%, and global output by 12%. The latter occurs, via the stimulation of production of new talents in Africa, Latin and Central America. Finally, the results indicate that Bosman rule decreases the inequality across the national teams by 70%.

Although the football labor market is a specific market, the results of this analysis provide an upper bound on the migration response to inequality for the global economy. If developed countries relax restrictions on highly-skilled immigration, the model developed in this study predicts greater inequality on a development level (as measured by GDP per capita), lower inequality in income per natural¹¹, and higher efficiency on a global level. The latter effect would be consequence of an increasing of production of talents by less developed regions due to stronger incentives to invest in education after the removal of quotas.

For further research, it would be interesting in investigating the evolution of inequality within European leagues. Presumably, it is expected to be increased since the richest clubs had now access to a larger pool of talents than before the Bosman rule. However, it is ambiguous as to whether the inequality within a league spurs productivity or no. Moreover, it is also crucial to explore whether a more unequal league is more or less appealing to the public than a more egalitarian league, given that other sports (like NBA league in Basketball) seek to make their leagues more equal and, thus, make games more exciting.

Acknowledgements

I thank the anonymous referee, the Associate Editor and the Editor in charge of this paper for constructive and helpful suggestions. Comments from Josh Angrist, Simone Bertoli, Raouf Boucekkin, Clemont de Chaisemartin, Frédéric Docquier, Mirco Draca, Chiara Farronato, Oded Galor, Sergey Mityakov, Ayan Orujov, Caglar Ozden, Chris Parsons, Marco Pelliccia, Giovanni Perri, Hillel Rapoport, Santos Silva, Kevin Tsui, Jerome Valette, Vincenzo Verardi, Fabian Waldinger and Jeffrey Wooldridge were also appreciated.

References

- Arellano, M. and Bond, S. (1991). "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations ", *Review of Economic Studies*, vol.58(2), pp. 277-97.
- Azariadis, C. and Drazen, A. (1990). "Threshold externalities in economic development ", *Quarterly Journal of Economics*, vol. 105, pp. 501-526.

¹¹Income per natural is fundamentally different from Gross National Income per capita. The former includes income to all people living outside their country of birth regardless of how long, while the latter includes only income to nationals of a country who reside temporarily abroad for less than one year (see Clemens and Pritchett, 2008).

- Beine, M., Docquier, F., Ozden, C. (2011). "Diasporas ", *Journal of Development Economics*, vol. 95, pp. 30-41.
- Beine, M., Docquier, F., Rapoport, H. (2008). "Brain drain and human capital formation in developing countries: winners and losers ", *Economic Journal*, vol. 118, pp. 631-652.
- Beine, M. and Parsons, C. (2015). "Climate factors as determinants of international migration ", *Scandinavian Journal Economics* vol.117(2), pp 723-767.
- Berlinschi, R., Swinnen, J. Schokkaert, J. (2013). "When Drains and Gains coincide: Migration and International Football Performance", *Labour economics*, vol.21, pp.1-14.
- Borjas, G. J. (2003). "The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market ", *Quarterly Journal of Economics* ,vol. 118(4),pp. 1335-1374.
- Canaday, N. and Tamura, R. (2009). "White discrimination in provision of black education: Plantations and towns ", *Journal of Economic Dynamics and Control*, vol.33, pp.1490-1530.
- Card, D. (2009), "Immigration and Inequality", *American Economic Review*, vol. 99(2), pp. 1-21.
- Caselli, F., Esquivel, G., Lefort, F. (1996). "Reopening the convergence debate: a new look at cross-country growth Empirics ", *Journal of Economic Growth*, vol.1, pp.363-389.
- Clair, G., Gaullier, G., Mayer, Th. , Zignago, S. (2004). "A note on CEPII's distances measures ", Explanatory note. CEPII, Paris.
- Clemens, M. and Pritchett, L. (2008), "Income Per Natural: Measuring Development for People rather than Places", *Population and Development Review* 34(3), pp. 395-434.
- de la Croix, D. and Docquier, F. (2012). "Do brain and poverty result from coordination failures?", *Journal of Economic Growth*, vol.17(1) pp.1-26.
- de Palma, A. and Kilani, K. (2007). "Invariance of conditional maximum utility ", *Journal of Economic Theory*, pp.137-146.
- Docquier, F., Lowell, L. and Marfouk A. (2009). "A gendered assessment of highly skilled emigration ", *Population and Development Review*, vol. 35(2), pp.297-322.
- Docquier, F. and Rapoport, H., (2012). "Globalization, brain drain and development ", *Journal of Economic Literature*, vol. 50 (3), pp. 681-730.
- Docquier, Machado J., Sekkat, K. (2015). Efficiency gains from liberalizing labor mobility ", *Scandinavian Journal of Economics* (2015), vol. 117 (2), pp. 303-346.
- Docquier, F., Lodigiani, E. Rapoport, H. Schiff, M. (2016). Emigration and democracy ", *Emigration and democracy* ", *Journal of Development Economics*, vol.120, pp.209-223.
- Ericson, T. (2000). "The Bosman case: Effects of the abolition of the transfer fee", *Journal of Sports Economics*, vol. 1 (3), pp. 203-218.
- Feyrer, J. (2009). "Trade and Income-Exploiting time series in Geography ", *NBER*, No14910.

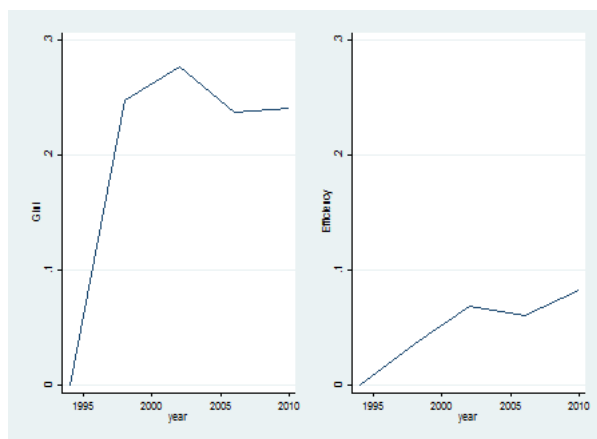
- Frick, B. (2009). "Globalization and factor mobility: the impact of the 'Bosman ruling' on player migration in professional soccer", *Journal of Sports Economics*, vol. 10 (1), pp. 88-106.
- Grogger, J. and Hanson, G. (2011). "Income maximization and the selection and sorting of international migrants ", *Journal of Development Economics*, vol.95, pp.42-57.
- Kennan J. and Walker, J. (2011). "The effect of expected Income on Individual migration decisions ", *Econometrica*, vol.79, pp.211-251.
- Kessenne, S. (2009). "The Impact of Pooling and Sharing Broadcast Rights in Professional Team Sports ", *International Journal of Sport Finance, Fitness Information Technology*, vol.4(3), pp.211-218.
- Kleven, H., Landais, C., Saez, E. (2013). "Taxation and International Migration of Superstars: Evidence from the European football market. " *American Economic Review*, vol.103(5), pp.1892-1924
- Lucas, R.E. (1988). "On the mechanics of economic development", *Journal of Monetary Economics*, vol 22, pp.3-42.
- McFadden, D. L. (1984), *Econometric Analysis of Qualitative Response Models*, in Z. Griliches and M. Intriligator (eds.), *Handbook of Econometrics*, Vol. 2, North-Holland, Amsterdam.
- Milanovic, B. (2005). "Globalization and Goals: Does Soccer show the Way? ", *Review of International Political Economy*, vol.88(1), pp.829-850.
- Mountford, A. (1997). "Can a brain drain be good for growth in the source economy?", *Journal of Development Economics*, vol 53, pp.287-303.
- Nickell, S. (1981). "Biases in Dynamic Models with Fixed effects ", *Econometrica*, vol.49(6), pp.1417-1426.
- Radoman, M. (2017). "Labor Market Implications of Institutional Changes in European Football. The Bosman Ruling and Its Effect on Productivity and Career Duration of Players ", *Journal of Sport Economics*, vol.8(7), pp.651-672.
- Rosen, S. (1981). "The economics of superstars ", *American Economic Review*, vol.71 (5), pp. 845-858.
- Santos Silva, J.M.C. and Tenreyro, S. (2006). "The Log of Gravity", *The Review of Economics and Statistics*, vol.88(4), pp. 641-658.
- Stark-Zimmermann, B. (2004). Retrieved from URL <http://www.national-football-teams.com>.
- Tamura, R. (1992). "Efficient Equilibrium Convergence: Heterogeneity and Growth ", *Journal of Economic Theory*, vol.58, pp.355-376.
- Tamura, R. (1994). "Fertility, human capital and the Wealth of Families ", *Economic Theory*, vol.4, pp.593-603.
- Tenreyro, S. (2007). "On the trade impact of nominal exchange rate volatility ", *Journal of Development Economics*, vol.82(2) pp.485-508.

Theoharides, C. (2014). "Manila to Malaysia, Quezon to Qatar: International Migration and the Effects of Origin-Country Human Capital, mimeo, University of Michigan, Ann Arbor.

Appendices

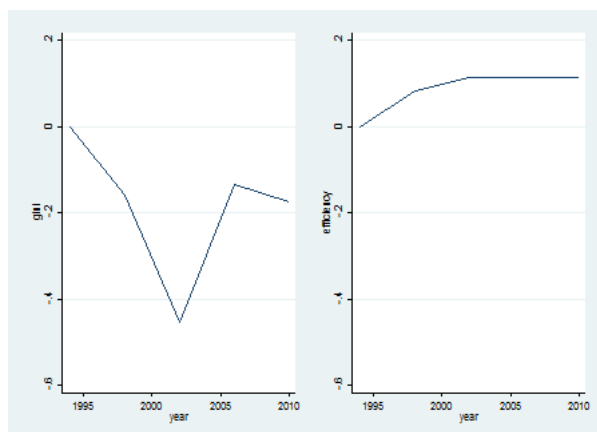
A Figures

Figure 1: Effect of Bosman rule on Inequality and Efficiency among European Leagues



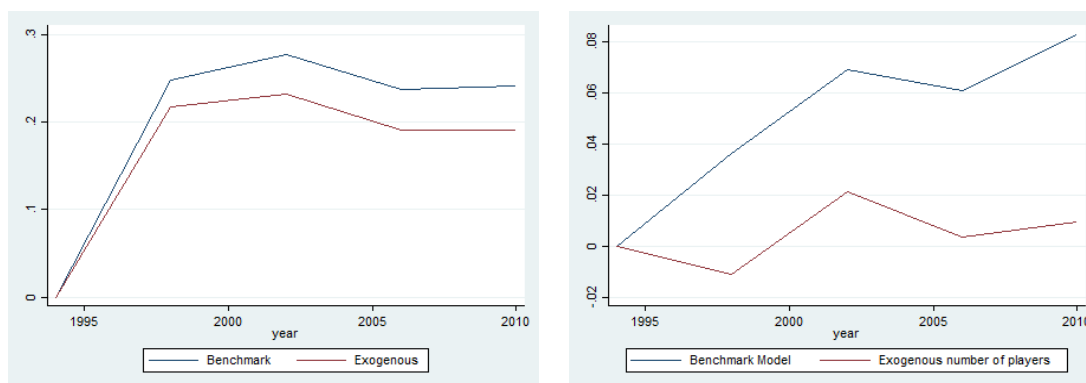
Note: The figure describes the inequality and efficiency across the years 1994 to 2010. GINI index and the leagues' score of Bosman trajectory are compared to those of the no-Bosman trajectory.

Figure 2: Effect of Bosman rule on Inequality and Efficiency among National Teams



The figure describes the national teams' inequality and efficiency across the years 1994 to 2010. The GINI index and the leagues' score of Bosman distribution are compared to those of the no-Bosman distribution.

Figure 3: Inequality and Efficiency for European Leagues

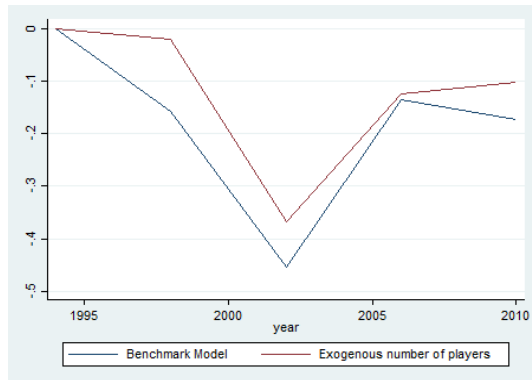


figure(3a)

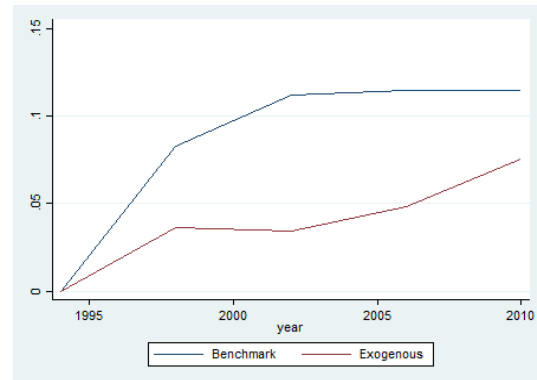
figure(3b)

Note: Figure 3a illustrates the inequality of European leagues under two scenarios. Scenario 1: benchmark model. Scenario 2: benchmark model with exogenous number of talented players. The GINI index of Bosman distribution is compared to that of the no-Bosman rule distribution. Figure 3b illustrates the efficiency of European leagues under two scenarios. Scenario 1: benchmark model. Scenario 2 : benchmark model with exogenous number of talents. Leagues' score of Bosman distribution is compared to that of the no-Bosman rule distribution. .

Figure 4: Inequality and Efficiency of National Teams



figure(4a)



figure(4b)

Note: Figure 4a illustrates the national teams' inequality under two scenarios. Scenario 1: benchmark model. Scenario 2: benchmark model with exogenous number of talented players. The Gini index of Bosman distribution is compared to that of the no-Bosman rule distribution. Figure 4b illustrates the European leagues' efficiency under two scenarios. Scenario 1: benchmark model. Scenario 2: benchmark model with exogenous number of talented players. National score of Bosman distribution is compared to that of the no-Bosman rule distribution.

B Tables

Table 1: Determinants of migration in soccer

Model :	Dependent Variable: Migrants/stayers ratio(logs)				
	EU x EU sample			All to EU sample	
	(1)	(2)	(3)	(4)	(5)
Colonial links	-0.023 (0.290)	0.272 (0.438)	0.273 (0.319)	-0.219 (0.379)	-0.223 (0.379)
Linguistic proximity	-0.153 (0.191)	-1.57 (0.200)	-0.194 (0.144)	-0.690 (0.400)	-0.688 (0.40)
Leagues' score ratio(logs)	0.427 ^a (0.152)	0.431 ^b (0.160)	1.318 ^b (0.634)	0.491 ^a (0.132)	0.444 ^a (0.127)
Distance (logs)	-2.138 ^a (0.138)	-2.105 ^a (0.151)	-2.158 ^a (0.106)	-3.691 ^a (0.245)	-3.691 ^a (0.245)
wages ratio (logs)	0.886 ^b (0.433)	0.927 ^c (0.520)	0.8200 ^c (0.52)	1.216 ^b (0.579)	1.277 ^b (0.619)
Bosman dummy	0.411 ^b (0.157)	0.454 ^b (0.220)	0.552 ^b (0.177)	1.034 ^a (0.225)	
Taxation		-0.407 (0.432)	-0.152 (0.437)	-0.364 (0.621)	-0.410 (0.621)
polity	0.052 (0.066)	0.053 (0.066)	0.071 (0.073)	2.274 ^a (0.658)	2.250 ^a (0.658)
<i>Bosman_{EU}</i>					0.627 ^a (0.224)
<i>Bosman_{NONEU}</i>					0.881 ^a (0.258)
Constant	13.11 ^a (3.31)	11.53 ^a (2.773)	10.598 ^a (3.905)	-4.111 (6.244)	-5.550 (6.28)
Destination Fixed Effects	Yes	Yes	Yes	Yes	Yes
Origin Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	No	No	No	No	No
Bspline	Yes	Yes	Yes	Yes	Yes
Years	1978-2010	1978-2010	1978-2010	1978-2010	1978-2010
Method	Poisson	Poisson	IV-Poisson	Poisson	Poisson
N	3154	2656	2191	4340	4340
R ²	0.55	0.52	0.62	0.61	0.61
P-value of Wald test for exogeneity			0.5943		
P-value of Wald test for different coefficients					0.278

Note: Standard errors clustered by country origin and destination in parentheses ^a, ^b and ^c denoting significance at the 1%, 5% and 10% levels respectively. The dependent variable in all the columns is Migrants/stayers ratio(logs). In Column 1, Eq.9 is estimated using the method of PPML for the sample 1978-2010 for European Countries to 21 European destination countries. Column 2 uses PPML method for the estimation of Eq.9 for the sample 1978-2010 for European countries including the variable of taxation(the number of destination countries falls to 14). Column 3 estimates the Eq.9 for the sample 1978-2010 using IV-poisson. Column 4 estimates Eq.9 using PPML including All to EU sample. Column 5 estimates Eq.9 using PPML including All to EU sample splitting the Bosman dummy to *Bosman_{EU}* *Bosman_{NONEU}*

Table 2: Estimation of the production function of leagues

Model :	(1)	(2)	(3)	(4)	(5)
	log league's score	log league's score	log league's score	log league's score	log league's score
number of talents(logs)	0.121 ^b (0.050)	0.110 ^a (0.033)	0.467 ^a (0.162)	0.611 ^a (0.166)	0.590 ^a (0.025)
Constant	-0.195 (0.191)	-0.508 ^a (0.183)	0.785 (0.170)	1.614 ^a (0.580)	1.420 ^a (0.146)
Country fixed Effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	185	254	140	185	183
R-square	0.82	0.85	0.71	0.70	0.68
Method	OLS	OLS	IV	IV	IV
F-test of first regression			28.59	13.07	14.10
Cragg-Donald test			28.59	20.1	22.05
Hausman test			10.2	6.32	10.2
p-value			0.0016	0.016	0.0046
Cumby-Huizinga test for autocorrelation			0.32	0.84	
Hansen test			0.3	0.42	

Standard errors clustered by country in parentheses ^a, ^b and ^c denoting significance at the 1%, 5% and 10% levels respectively. Talent is defined the soccer player who has at least three appearances with his national team. Column 1 estimates Eq.10 using OLS method for the sample UEFA countries without the Eastern European countries. Column 2 estimates Eq.10 using OLS method for the sample including all the European countries. Column 3 estimates Eq.10 using IV-method for the sample UEFA countries without Eastern European countries. Column 4 estimates Eq.10 using IV-method for the sample all the European leagues. The first two lags are used as instruments for the variable number of migrants (logs). Column 5 estimates Eq.10 using IV-method for the sample UEFA countries without Eastern European countries using an exogenous instrument following the methodology of Feyrer (2009).

Table 3: Estimation of the production function of national teams

	(1)	(2)	(3)
	national score (logs)	national score (logs)	national score (logs)
Total level of skills (logs)	0.084 ^b (0.039)	0.733 ^b (0.384)	0.150 ^a (0.04)
constant	-0.601 ^a (0.0636)	-0.506 ^a (0.212)	-0.420 (0.110)
Country fixed Effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	309	186	304
R-square	0.75	0.75	0.63
F-test of first regression		11.5	50
Cragg-Donald test		39.5	121
Cumby-Huizinga test for autocorrelation(p-value)		0.10	
Hansen test		0.90	

Notes: Standard errors clustered by country in parentheses ^a, ^b and ^c denoting significance at the 1%, 5% and 10% levels respectively. Column 1 and 3 estimates Eq. 12 using OLS and IV method. The first two lags are used as instruments for the variable Total level of skills (logs). Column 2 estimates Eq. 12 using OLS by disentangle the total level of skills into Average skills for talents and Number of talents. Last column estimates Eq. 12 using IV method with instrument following the methodology of Feyrer (2009).

Table 4: Training for soccer players(Dependent :Variation in Number of talented players)

	(1)	(2)	(3)
Lagged number of talents(logs)	-0.690 ^a (0.056)	-0.841 ^c (0.474)	
Lagged variation in Number of talents (logs)			0.435 ^a (0.128)
Bosman x Europe	-0.020 (0.330)	0.125 (0.150)	0.164 (0.294)
Bosman x Africa	1.164 ^a (0.350)	1.04 ^a (0.611)	0.66 ^b (0.324)
Bosman x Asia	0.131 (0.218)	0.368 (0.298)	0.326 (0.355)
Bosman x Central America	1.033 ^a (0.337)	1.16 ^a (0.570)	0.592 ^a (0.294)
Bosman x Latin America	1.140 ^a (0.330)	1.15 ^a (0.560)	0.706 ^a (0.225)
Bosman x B.A	-0.208 (0.185)	-0.030 (0.208)	-0.063 (0.209)
logtrain	0.176 (0.183)	0.437 (0.514)	0.173 ^a (0.043)
polity	0.001 (0.02)	0.005 (0.06)	0.014 (0.0203)
Country fixed Effects	Yes	Yes	No
Year fixed effects	Yes	Yes	Yes
Observations	360	260	306
R-square	0.59	0.49	
F-first reg,		17	
Cragg-Donald		22.3	
AR(1)p-value			0.000
AR(2)p-value			0.562
Number of groups			51
Number of Instruments			30
Hansen test		0.11	0.131

Notes: Standard errors are clustered by countries in parentheses ^a, ^b and ^c denoting significance at the 1%, 5% and 10% levels respectively. Column 1 estimates Eq.16 using OLS method. Column 2 estimates Eq.16 using IV method. Column 3 estimates Eq.16 using Arellano Bond method. As instruments are used the first two lags of talented players. The number of instruments is 30 and number of groups is 51.

Table 5: Calibrated values

Coefficients	values
λ	0.77
γ	0.53
δ	0.7
φ_{Africa}	0.66
$\varphi_{North-CentralAmerica}$	0.59
$\varphi_{SouthAmerica}$	0.71
$\varphi_{Europe-Asia-BA}$	0